A Hybrid ICA-SVM Approach to Continuous Phase Modulation Recognition

David Boutte, *Member, IEEE*, and Balu Santhanam, *Senior Member, IEEE*

***Abstract*—Automatic modulation recognition is a topic of interest in many fields including signal surveillance, multi-user detection and radio frequency spectrum monitoring. In this paper, we present an algorithm for recognition of different types of continuous phase modulation signals that uses a combination of features extracted through cyclic spectral analysis and an ICA-SVM hybrid recognition system. Simulation results demonstrate the ability of the algorithm to correctly identify modulation types over a wide range of SNR scenarios. The effects of pulse shaping and partial response waveforms are also investigated.**

***Index Terms*—Automatic modulation recognition, ICA, SVM.**

I. Introduction

AUTOMATIC modulation identification plays an important role in signal surveillance and frequency spectrum monitoring. The role of modulation recognition is to determine the modulation type and associated parameters from a received signal. A class of existing modulation recognition approaches make use of instantaneous signal parameters [1]–[4]. However, these algorithms fail to deal with many types of *continuous phase modulation* (CPM) [5]. CPM, which is widely used in wireless communications, uses a constant envelope waveform and phase continuity to achieve high spectral efficiency. Classification of CPM signals was studied in [6], however, several limitations of the feature extraction proposed there caused increased misclassification between similar types of CPM schemes. An algorithm based on *independent component analysis* (ICA) and *support vector machines* (SVMs) which can classify several types of CPM signals is proposed.

Most digital modulation schemes exhibit cyclostationary behavior [7], described by the spectral correlation density function (SCD). The estimation of this function is based on the smoothed cyclic periodogram, defined as

$$S^{\alpha}_{XX_T} = \frac{1}{T}\left\langle X_T\left(n, f+\frac{\alpha}{2}\right) X^*_T\left(n, f-\frac{\alpha}{2}\right)\right\rangle \quad (1)$$

Manuscript received September 04, 2008; revised January 18, 2009. Current version published March 26, 2009. This work was supported by the Miratek Corporation. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Weifeng Su.

The authors are with the Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131-001 USA (e-mail: dboutte@ece.unm.edu; bsanthan@ece.unm.edu).

Color versions of one or more figures in this paper are available online at http://ieeexplore.ieee.ogr.

Digital Object Identifier 10.1109/LSP.2009.2016444

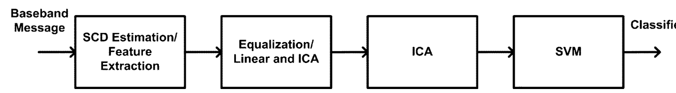


Fig. 1. System Block Diagram. The feature extraction block is composed of SCD estimation and parameterization. The decision tree is generated using a SVM. Combined with the ICA this allows non-Gaussian signal statistics to be fully exploited.

where $\langle\cdot\rangle$ denotes the time average operation. $X_T(n, f+(\alpha/2))$ is the frequency shifted complex demodulate of $x(n)$ computed as

$$X_T(n,f) = \sum_{r=-\frac{M}{2}}^{\frac{M}{2}} a(r)x(n-r)e^{-i2\pi f(n-r)T} \quad (2)$$

where $a(r)$ is the Hamming spectral window function. Using the estimate of the SCD, the modulation of interest's features can be extracted and a classifier developed using techniques in ICA and SVMs.

Independent component analysis (ICA) is a statistical method for separating a multivariate signal into additive subcomponents supposing the mutual independence of the source signals [8]. This technique is a subclass of source separation algorithms, which has been widely applied in blind source separation. It has also been directly applied to the modulation recognition problem in [9]. However, in this paper; the ICA algorithm is used as a post processing tool that will remove statistical dependence between the features by exploiting their non-Gaussianity with the aid of higher order statistics.

The field of statistical learning has many tools which can be applied to the modulation recognition problem. *Support vector machines* (SVMs) offer a complex decision making process without much of the demands on SNR and signal frame length. SVMs provide an efficient means of classification without placing assumptions on signal statistics or on SNRs. The standard binary SVM classifier as described in [10] can be simply extended to a multi-class classifier for use in this case. By appropriately weighting the features with respect to importance, problems of misclassifying similar modulations can be mitigated.

In this paper, we propose the use of cyclostationary statistical analysis combined with ICA-SVM to form a hybrid approach to the problem of continuous phase modulation recognition. The optimal solution would require the joint optimization of the ICA weights and contrast function together with SVM support vectors and kernel. However, due to the number of parameters involved in each operation the joint optimization problem becomes intractable. A suboptimal solution is presented in this paper which independently optimizes the ICA and SVM modules. By combining these methods together into a hybrid approach, an efficient and robust algorithm is developed that exhibits superior performance over a wide range of SNR values.

A. Algorithm

The proposed algorithm consists of three parts described in Fig. 1 first the SCD is estimated and then parameterized in terms of statistical features [11]. These cyclic features are the standard deviation, the autocorrelation and the kurtosis of the SCD. Second, the feature vectors are passed to an ICA using denoising techniques [12]. Third, the ICA transformed features are passed to a composite SVM using feature weighting to determine the appropriate classifier.

The ICA is implemented using a FASTICA algorithm with QR preprocessing [13]. The ICA processing is used in two ways. First, it is a way to force statistical independence on the features by separating them into independent sources. This is done by exploiting the non-Gaussian feature distributions. In cases where there is minimal non-Gaussianity present, the ICA will not offer an improvement in feature distribution. Empirical trials have shown it is advantageous to force statistical independence on the features so that the SVM can more easily construct a reliable classifier with SNR lower than 14 dB. The second use of the ICA processing is as a denoising tool. In this case, it functions similar to a modified whitening operation where the effects of AWGN can be suppressed and the probability of recognition improved.

The SVM module is implemented as a *combined one versus all* algorithm [10] where the feature vectors are adaptively weighted in importance.

Empirical trials have shown that individual features may have more relevance in classifying one modulation as opposed to another. To exploit this; once an initial classifier is constructed the feature weights are adjusted and the SVM is rerun on the weighted features, or bootstrapped into the existing classifier. A complete discussion of the SVM risk minimization principle can be found in [14]. The effect of this weighting is the modification of the support vectors so that the distance in the kernel transformed space is maximized. This produces an iteratively converging classifier. A reliable decision is reached after about 20 iterations of this procedure.

The kernel used here is in the family of radial basis functions. It is specifically, a truncated Taylor expansion of a radial basis function. It was chosen to combine the simplicity inherent in polynomial kernels and the robustness of Gaussian kernels [14]. The statistical independence introduced by the ICA manifests itself as L_2 distance in the kernel transformed space of the SVM.

The use of ICA-SVMs has been proposed before by [15]. However there are several differences in the algorithm proposed here. The use of ICA in [15] is as a method of extracting a feature set. In this proposed algorithm, the features are extracted from the cyclic spectrum and the ICA is a tool used to refine the features and reduce the classifier error present in the SVM. Empirical trials have shown a connection between SVM margin and input negentropy, which justifies feature preprocessing. This algorithm was applied to two types of CPM; varied pulse shape CPM and varied memory CPM. These results are presented in the following sections.

Modulation Type	Normalized Kurtosis
5REC-CPM KUR Feature	2.23
MSK KUR Feature	2.4
GMSK KUR Feature	2.5
3REC-CPM AC Feature	2.51
CPFSK STD Feature	2.65
1RAC-CPM AC Feature	2.91
1REC-CPM AC Feature	3

Fig. 2. Normalized kurtosis values of the priority weighted feature of several CPM schemes. These features vary from modulation type to modulation type however they have the greatest impact upon classifier construction the SVM. Values greater than one indicate a super-Gaussian distribution.

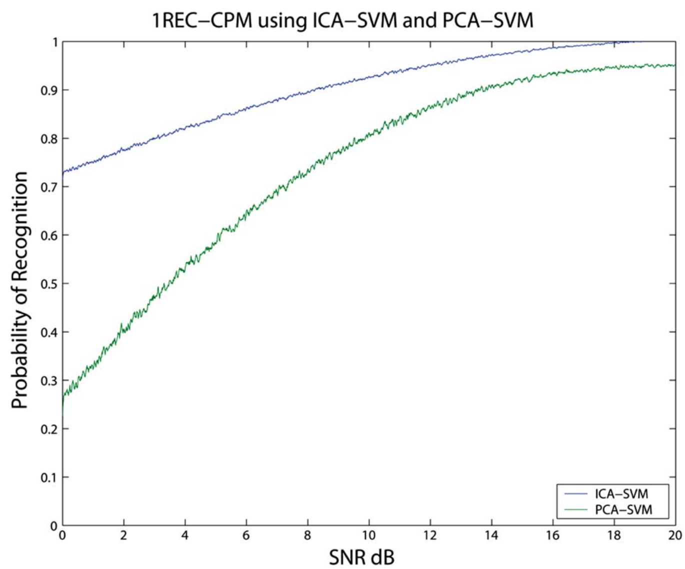


Fig. 3. Probability of Recognition of CP-FSK using both PCA and ICA processing using 1000 experiments. The improvement in performance with the ICA is indicative of the non-Gaussian nature of the features. This benefit is manifested as more reliable classifiers in the SVM at higher noise powers.

II. Results and Analysis

An algorithm as outlined in Fig. 1 was implemented using MATLAB. A set of varying memory rectangular (REC) and raised cosine (RAC) pulse shaped CPM signals were designed and studied at baseband. Since the ICA chiefly exploits non-Gaussianity present in the input sources, a measure of this non-Gaussianity is described in Fig. 2. Non-Gaussianity is characterized by comparing the normalized kurtosis of a feature with unity. The distribution is either sub or super-Gaussian if the normalized kurtosis is less than or greater than one. The types of CPM studied here have features that have super-Gaussian distributions. The features that are present in Fig. 2 are those features that are weighted in importance inside the SVM. The effect of this is shown in Fig. 3 where the ICA is replaced with a similar PCA algorithm. The performance of the algorithm using a PCA is significantly degraded in principle because of the non-Gaussian distribution of the feature vectors.

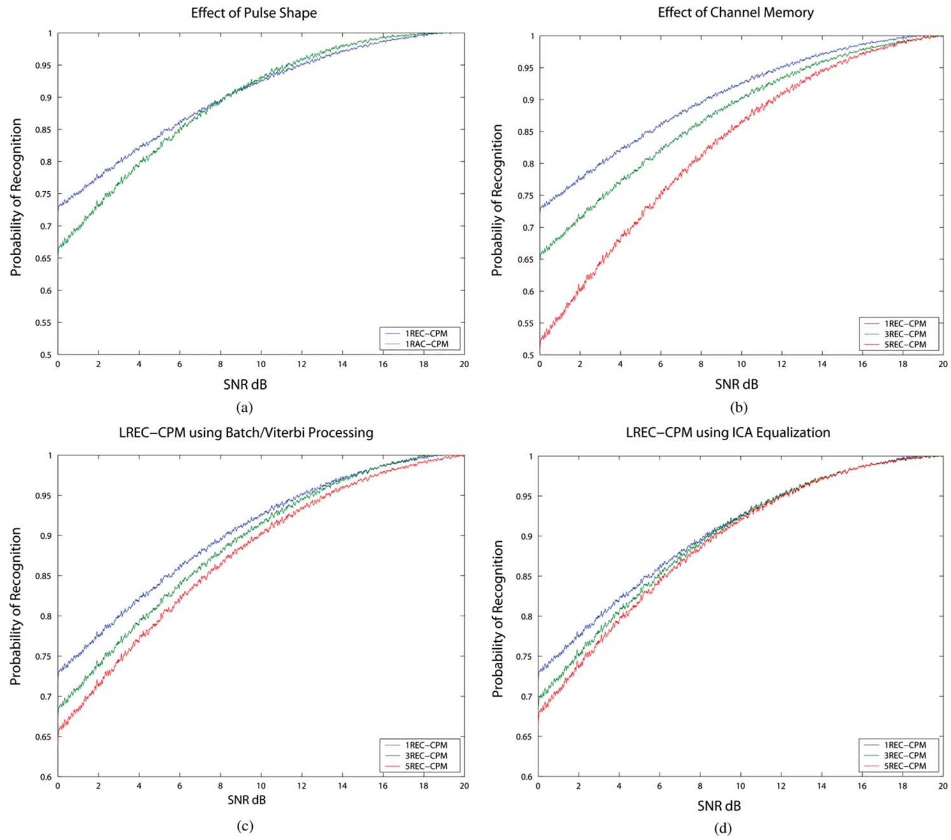


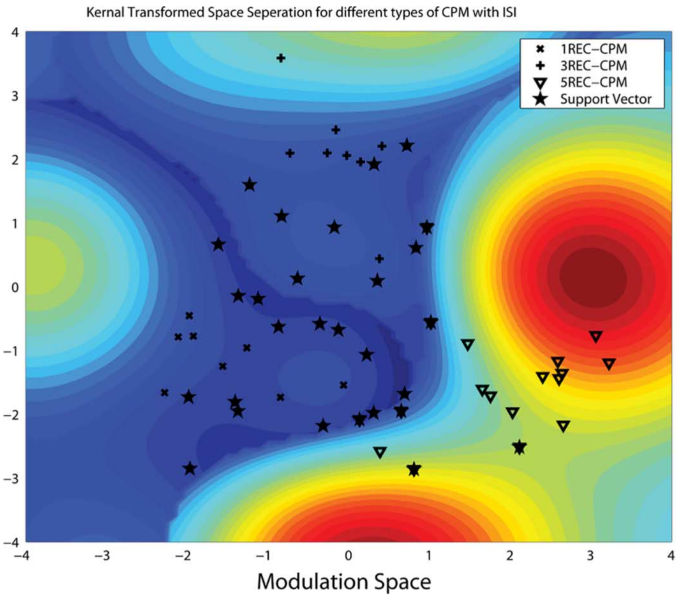
Fig. 4. Probability of recognition for CPM using different amounts of channel memory and pulse shapes. Changes in the pulse shape seem to have little impact upon recognition, while changes in the memory have a more significant impact. This is primarily due to increasing inter symbol interference in higher memory cases. When Viterbi equalization is used the performance is improved. In contrast an ICA based equalizer has a slightly better performance with higher memory modulations. The small separation error between the modulations is attributed to noise effects present in the equalizers. (a) Varying pulse shape, (b) Varying memory, (c) Using Viterbi equalization, (d) Using ICA equalization.

The effect of a particular pulse shape does not have significant impact upon the algorithm's performance as seen in Fig. 4(a). The algorithm's probability of recognition was studied using different types of pulse shaping functions, specifically rectangular and raised cosine. These changes in pulse shape appear to have no significant effect upon the probability of recognition. The impact of pulse shaping functions is seen primarily in the power spectral density (PSD), or cycle frequency equal to zero in the SCD. The bandwidth reduction offered by a particular pulse shaping function has little impact beyond the PSD in altering specific cycle frequencies or changing the shape of the SCD. This limited impact is manifested in Fig. 4(a) where two separate pulse shapes yield similar performance.

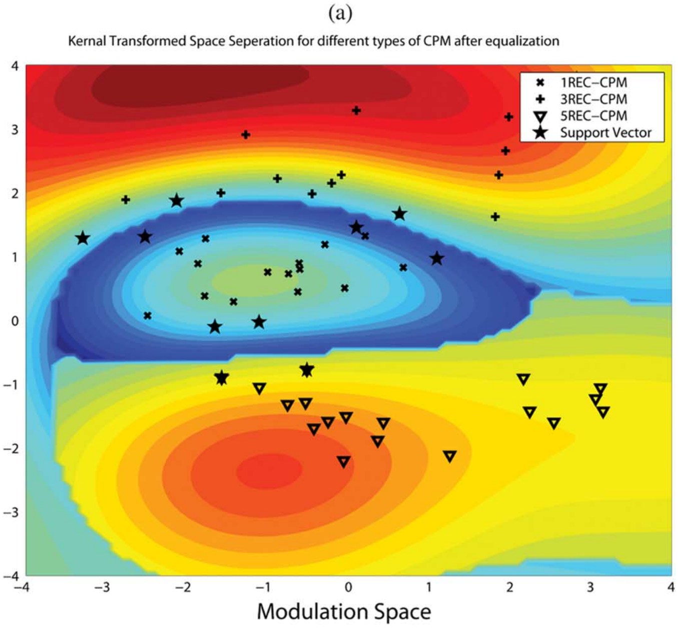
However, altering the channel memory, i.e., the number of bits effected, does have a significant impact upon performance. Fig. 4(b) shows the effect channel memory has on system performance. For a set rectangular pulse shape the value of L, channel memory, was varied and the effects upon performance studied. As shown, as the value of L is increased the probability of recognition is decreased. This is primarily caused by the presence of *inter symbol interference* (ISI) and as L is increased; the weighted feature's non-Gaussianity is decreasing. This causes the output from the ICA to be suboptimal and further degrades performance. This can be improved upon however by changing the feature processing algorithm. Instead of using per bit processing, switching to batch processing with equalization is a way to suppress ISI. This is illustrated in Fig. 5 using similar types of CPM. Instead of performance plots, these are contour maps of the kernel transformed space inside the SVM. As seen in Fig. 5(a), the separation between the modulations using the vanilla algorithm is very small in the presence of ISI. This margin of separation is increased, when equalization is employed as seen in Fig. 5(b). The increased margin translates directly into an increase in probability of recognition as was shown in Fig. 4.

III. Conclusion

We have presented in this letter an automatic modulation recognition algorithm for CPM signals. Exploiting cyclostationarity of the modulations and the non-Gaussian statistical characteristics of the modulation allows the extraction of features with minimal mutual information. The empirical nature of the SVM allows recognition without unrealistic assumptions of SNRs or training duration. Simulation results demonstrate a robust recognition performance over a wide range of SNRs, different pulse shaping functions and channel memory.



(a)



(b)

Fig. 5. Kernel transformed spaces for different types of CPM that contain ISI. Notice the large number of support vectors needed when ISI is present and the small separation between modulations. This lowers the probability of recognition for the various types of modulations. After equalization, the margin of separation between the classes is improved.

REFERENCES

[1] E. E. Azzouz and A. K. Nandi, *Automatic Modulation Recognition of Communication Signals*. Dordrecht, The Netherlands: Kluwer, 1996, 3300 AA Dordrect.

[2] M. L. D. Wong and A. K. Nandi, "Automatic digital modulation recognition using spectral and statistical features with multi-layer perceptrons," in *Proc. Sixth Int. Symp. Signal Processing and Its Applications*, 2001, vol. 2, pp. 390–393.

[3] A. Swami and B. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun.*, vol. 48, pp. 416–429, Mar. 2000.

[4] W. A. Gardner and S. Venkataraman, *Performance of Optimum and Adaptive Frequency-Shift Filters for Cochannel Interference and Fading*. San Jose, CA: Maple Press, 1990.

[5] T. Aulin and C. Sundberg, "Continuous phase modulation-Part I: Full response signaling," *IEEE Trans. Commun.*, vol. 29, pp. 187–195, 1981.

[6] F. Liedtke, "Adaptive procedure for automatic modulation recognition," *J. Telecommun. Inform. Technol.*, vol. 4, pp. 91–97, 2004.

[7] W. A. Gardner, "Two alternative philosophies for estimation of the parameters of time-series," *IEEE Trans. Inform. Theory*, vol. 37, pp. 216–218, Jan. 1991.

[8] A. Hyvarinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Netw.*, vol. 13, pp. 411–430, Feb. 2000.

[9] L. Hong, "Classification of bpsk and qpsk signals in fading environment using the ICA technique," in *Proc. Thirty-Seventh Southeastern Symp. System Theory*, 2005, vol. 1, pp. 471–475.

[10] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[11] D. Boutte and B. Santhanam, "A feature weighted hybrid ICA-SVM approach to automatic modulation recognition," in *Proc. IEEE Digital Signal Processing Workshop*, 2009, pp. 399–403.

[12] D. Zhang and P. K. Rajan, "Supergaussian data denoising by semi-ICA estimation," in *Proc. Thirty-Seventh Southeastern Symp. System Theory*, 2005, vol. 3, pp. 485–490.

[13] M. Gupta and B. Santhanam, "A general approach towards blind multiuser detection using higher-order statistics," *Proc. IEEE WCNC*, vol. 3, pp. 1585–1590, 2006.

[14] V. N. Vapnik, *Statistical Learning Theory*. New York, New York: Wiley, 1998.

[15] Y. Qi, D. Doermann, and D. DeMethon, "Hybrid independent component analysis and support vector machinelearning scheme for face detection," in *Proc. IEEE Int. Conf. Acoustics Speech and Signal Processing*, 2001, vol. 2, pp. 1481–1484.

[16] C. M. Spooner and W. A. Gardner, "Estimation of cyclic polyspectra," in *Proc. Twenty-Fifth Asilomar Conf. Signals and Systems*, 1991, vol. 2, pp. 370–376.

[17] D. Vucic and M. M. Obrodavic, "Cyclic spectral analysis of phase-incoherent fsk signal by matrix-based stochastic method," in *Proc. 5th Int. Conf. Telecommunications in Modern Satellite Cable and Broadcasting Service*, 2001, vol. 1, pp. 267–270.